



#2

OIEP

## RAW SEQUENCE LISTING

DATE: 02/25/2002

PATENT APPLICATION: US/10/067,291

TIME: 11:54:03

Input Set : N:\Crf3\RULE60\10067291.txt

Output Set: N:\CRF3\02252002\J067291.raw

3 <110> APPLICANT: Takashima, Yoshiki  
 4 Mitsuda, Satoshi  
 6 <120> TITLE OF INVENTION: Protein capable of catalyzing transamination  
 7 stereoselectively, gene encoding said protein  
 8 and use thereof  
 10 <130> FILE REFERENCE: Q58251  
 C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/067,291  
 13 <141> CURRENT FILING DATE: 2002-02-07  
 15 <150> PRIOR APPLICATION NUMBER: 09/527,522  
 16 <151> PRIOR FILING DATE: 2000-03-16  
 19 <150> PRIOR APPLICATION NUMBER: JP 11/075511  
 20 <151> PRIOR FILING DATE: 1999-03-19  
 22 <150> PRIOR APPLICATION NUMBER: JP 11/088634  
 23 <151> PRIOR FILING DATE: 1999-03-30  
 25 <160> NUMBER OF SEQ ID NOS: 12  
 27 <170> SOFTWARE: PatentIn Ver. 2.1  
 29 <210> SEQ ID NO: 1  
 30 <211> LENGTH: 339  
 31 <212> TYPE: PRT  
 32 <213> ORGANISM: Mycobacterium aurum  
 34 <220> FEATURE:  
 35 <223> OTHER INFORMATION: SC-S423  
 37 <400> SEQUENCE: 1  
 38 Met Thr Ala Leu Ser Asp Leu Gly Thr Ser Asn Leu Val Ala Val Glu  
 39 1 5 10 15  
 41 Pro Gly Ala Ile Arg Glu Asp Thr Pro Ala Gly Ser Val Ile Gln Tyr  
 42 20 25 30  
 44 Ser Asp Tyr Glu Leu Asp Thr Ser Ser Pro Phe Ala Gly Gly Val Ala  
 45 35 40 45  
 47 Trp Ile Glu Gly Glu Tyr Leu Pro Ala Glu Glu Ala Lys Ile Ser Ile  
 48 50 55 60  
 50 Phe Asp Thr Gly Phe Gly His Ser Asp Leu Thr Tyr Thr Val Ala His  
 51 65 70 75 80  
 53 Val Trp His Gly Asn Ile Phe Arg Leu Gly Asp His Leu Asp Arg Leu  
 54 85 90 95  
 56 Leu Asp Gly Ala Ser Lys Leu Arg Leu Asp Ala Gly Tyr Ser Lys Asp  
 57 100 105 110  
 59 Glu Leu Ala Glu Ile Thr Lys Lys Cys Val Ser Met Ser Gln Leu Arg  
 60 115 120 125  
 62 Glu Ser Phe Val Asn Leu Thr Val Thr Arg Gly Tyr Gly Lys Arg Lys  
 63 130 135 140  
 65 Gly Glu Lys Asp Leu Ser Lys Leu Thr His Gln Val Tyr Ile Tyr Ala  
 66 145 150 155 160

ENTERED

## RAW SEQUENCE LISTING

DATE: 02/25/2002

PATENT APPLICATION: US/10/067,291

TIME: 11:54:03

Input Set : N:\Crif3\RULE60\10067291.txt

Output Set: N:\CRF3\02252002\J067291.raw

```

68 Ile Pro Tyr Leu Trp Ala Phe Pro Pro Ala Glu Gln Ile Phe Gly Thr
69           165           170           175
71 Thr Ala Ile Val Pro Arg His Val Arg Arg Ala Gly Arg Asn Thr Val
72           180           185           190
74 Asp Pro Thr Ile Lys Asn Tyr Gln Trp Gly Asp Leu Thr Ala Ala Ser
75           195           200           205
77 Phe Glu Ala Lys Asp Arg Gly Ala Arg Thr Ala Ile Leu Leu Asp Ser
78           210           215           220
80 Asp Asn Cys Val Ala Glu Gly Pro Gly Phe Asn Val Cys Ile Val Lys
81 225           230           235           240
83 Asp Gly Lys Leu Ala Ser Pro Ser Arg Asn Ala Leu Pro Gly Ile Thr
84           245           250           255
86 Arg Lys Thr Val Phe Glu Leu Ala Asp Gln Met Gly Ile Glu Ala Thr
87           260           265           270
89 Leu Arg Asp Val Thr Ser Arg Glu Leu Tyr Asp Ala Asp Glu Leu Met
90           275           280           285
92 Ala Val Thr Thr Ala Gly Gly Val Thr Pro Ile Asn Ser Leu Asp Gly
93           290           295           300
95 Glu Ala Val Gly Asn Gly Glu Pro Gly Pro Leu Thr Val Ala Ile Arg
96 305           310           315           320
98 Asp Arg Phe Trp Ala Leu Met Asp Glu Pro Gly Pro Leu Ile Glu Thr
99           325           330           335
101 Ile Glu Tyr
104 <210> SEQ ID NO: 2
105 <211> LENGTH: 1020
106 <212> TYPE: DNA
107 <213> ORGANISM: Mycobacterium aurum
109 <220> FEATURE:
110 <223> OTHER INFORMATION: SC-S423
112 <220> FEATURE:
113 <221> NAME/KEY: CDS
114 <222> LOCATION: (1)..(1017)
116 <400> SEQUENCE: 2
117 atg act gct ctt tca gac ctc ggc acc tcc aac ctg gtg gcc gtc gag 48
118 Met Thr Ala Leu Ser Asp Leu Gly Thr Ser Asn Leu Val Ala Val Glu
119 1 5 10 15
121 ccc ggc gcc atc cgc gag gac acc ccg gcc ggc tcg gtg atc cag tac 96
122 Pro Gly Ala Ile Arg Glu Asp Thr Pro Ala Gly Ser Val Ile Gln Tyr
123 20 25 30
125 agc gac tac gaa ctg gac acc tcc agc ccg ttc gcc ggc gtc gcc 144
126 Ser Asp Tyr Glu Leu Asp Thr Ser Ser Pro Phe Ala Gly Gly Val Ala
127 35 40 45
129 tgg atc gag ggc gaa tac ctg ccg gcc gaa gaa gcg aag atc tcc atc 192
130 Trp Ile Glu Gly Glu Tyr Leu Pro Ala Glu Glu Ala Lys Ile Ser Ile
131 50 55 60
133 ttc gac acc gga ttc ggt cat tcc gat ctg acc tac acc gtc gcg cat 240
134 Phe Asp Thr Gly Phe Gly His Ser Asp Leu Thr Tyr Thr Val Ala His
135 65 70 75 80
137 gta tgg cac ggc aac atc ttc cgg ctc ggc gac cac ctg gac cgg ttg 288

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/067,291

DATE: 02/25/2002

TIME: 11:54:03

Input Set : N:\Crf3\RULE60\10067291.txt

Output Set: N:\CRF3\02252002\J067291.raw

138	Val	Trp	His	Gly	Asn	Ile	Phe	Arg	Leu	Gly	Asp	His	Leu	Asp	Arg	Leu	
139					85					90					95		
141	ctc	gac	ggg	gcg	tcc	aag	ctg	cgc	ctg	gac	gcc	ggg	tac	agc	aag	gac	336
142	Leu	Asp	Gly	Ala	Ser	Lys	Leu	Arg	Leu	Asp	Ala	Gly	Tyr	Ser	Lys	Asp	
143				100					105					110			
145	gaa	ctg	gcc	gag	atc	acc	aag	aag	tgc	gtg	tcg	atg	tcg	cag	ctg	cgc	384
146	Glu	Leu	Ala	Glu	Ile	Thr	Lys	Lys	Cys	Val	Ser	Met	Ser	Gln	Leu	Arg	
147			115					120					125				
149	gaa	tcg	ttc	gtg	aat	ctg	acc	gtc	acc	cgg	gga	tac	gga	aag	cgc	aag	432
150	Glu	Ser	Phe	Val	Asn	Leu	Thr	Val	Thr	Arg	Gly	Tyr	Gly	Lys	Arg	Lys	
151		130					135					140					
153	ggc	gag	aag	gac	ctg	tcc	aag	ctc	acc	cat	cag	gtg	tac	atc	tac	gcc	480
154	Gly	Glu	Lys	Asp	Leu	Ser	Lys	Leu	Thr	His	Gln	Val	Tyr	Ile	Tyr	Ala	
155	145					150					155					160	
157	atc	ccg	tac	ctg	tgg	gcc	ttc	ccg	ccc	gcc	gag	cag	atc	ttc	ggc	acc	528
158	Ile	Pro	Tyr	Leu	Trp	Ala	Phe	Pro	Pro	Ala	Glu	Gln	Ile	Phe	Gly	Thr	
159				165					170					175			
161	acc	gcg	atc	gtg	ccg	cgc	cat	gtc	cgc	cgc	gcc	ggc	cgc	aac	acc	gtc	576
162	Thr	Ala	Ile	Val	Pro	Arg	His	Val	Arg	Arg	Ala	Gly	Arg	Asn	Thr	Val	
163				180					185					190			
165	gac	ccg	acc	atc	aag	aac	tac	cag	tgg	ggt	gat	ctc	acc	gca	gcc	agt	624
166	Asp	Pro	Thr	Ile	Lys	Asn	Tyr	Gln	Trp	Gly	Asp	Leu	Thr	Ala	Ala	Ser	
167			195					200					205				
169	ttc	gaa	gcc	aag	gac	cgt	ggt	gcg	cgc	acc	gcg	atc	ctg	ctc	gac	tcg	672
170	Phe	Glu	Ala	Lys	Asp	Arg	Gly	Ala	Arg	Thr	Ala	Ile	Leu	Leu	Asp	Ser	
171		210					215					220					
173	gac	aac	tgc	gtg	gcc	gaa	ggt	ccg	ggc	ttc	aac	gtg	tgc	atc	gtc	aag	720
174	Asp	Asn	Cys	Val	Ala	Glu	Gly	Pro	Gly	Phe	Asn	Val	Cys	Ile	Val	Lys	
175	225					230				235						240	
177	gac	ggc	aag	ctg	gcc	tcc	ccg	tcc	cgg	aac	gcg	ttg	ccg	ggc	atc	acc	768
178	Asp	Gly	Lys	Leu	Ala	Ser	Pro	Ser	Arg	Asn	Ala	Leu	Pro	Gly	Ile	Thr	
179				245					250					255			
181	cgt	aag	acg	gtg	ttc	gaa	ctg	gcc	gac	cag	atg	ggc	atc	gaa	gcc	acc	816
182	Arg	Lys	Thr	Val	Phe	Glu	Leu	Ala	Asp	Gln	Met	Gly	Ile	Glu	Ala	Thr	
183				260					265					270			
185	ctg	cgc	gac	gtc	acc	agc	cgt	gaa	ctc	tac	gac	gcc	gac	gag	ttg	atg	864
186	Leu	Arg	Asp	Val	Thr	Ser	Arg	Glu	Leu	Tyr	Asp	Ala	Asp	Glu	Leu	Met	
187			275					280					285				
189	gcg	gtc	acc	acc	gcg	ggc	ggg	gtc	aca	ccg	atc	aac	tcg	ctg	gat	ggc	912
190	Ala	Val	Thr	Thr	Ala	Gly	Gly	Val	Thr	Pro	Ile	Asn	Ser	Leu	Asp	Gly	
191		290					295					300					
193	gag	gcc	gtg	ggc	aac	ggc	gag	ccc	ggt	cca	ctg	acg	gtg	gcc	atc	cgc	960
194	Glu	Ala	Val	Gly	Asn	Gly	Glu	Pro	Gly	Pro	Leu	Thr	Val	Ala	Ile	Arg	
195	305					310				315						320	
197	gac	cgc	ttc	tgg	gcg	ctg	atg	gac	gag	ccg	ggc	ccg	ctg	atc	gaa	acg	1008
198	Asp	Arg	Phe	Trp	Ala	Leu	Met	Asp	Glu	Pro	Gly	Pro	Leu	Ile	Glu	Thr	
199				325					330					335			
201	atc	gaa	tac	tga													1020
202	Ile	Glu	Tyr														

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/067,291

DATE: 02/25/2002

TIME: 11:54:03

Input Set : N:\Crf3\RULE60\10067291.txt

Output Set: N:\CRF3\02252002\J067291.raw

```

205 <210> SEQ ID NO: 3
206 <211> LENGTH: 18
207 <212> TYPE: PRT
208 <213> ORGANISM: Mycobacterium aurum
210 <220> FEATURE:
211 <223> OTHER INFORMATION: SC-S423
213 <220> FEATURE:
214 <221> NAME/KEY: UNSURE
215 <222> LOCATION: (4)
217 <220> FEATURE:
218 <221> NAME/KEY: UNSURE
219 <222> LOCATION: (9)
221 <400> SEQUENCE: 3
W--> 222 Thr Ala Leu Xaa Asp Leu Gly Thr Xaa Asn Leu Val Ala Val Glu Pro
      223   1             5             10             15
      225 Gly Ala
      228 <210> SEQ ID NO: 4
      229 <211> LENGTH: 17
      230 <212> TYPE: PRT
      231 <213> ORGANISM: Mycobacterium aurum
      233 <220> FEATURE:
      234 <223> OTHER INFORMATION: SC-S423
      236 <400> SEQUENCE: 4
      237 Ile Ser Ile Phe Asp Thr Gly Phe Gly Ala Ser Asp Leu Thr Tyr Thr
      238   1             5             10             15
      240 Val
      243 <210> SEQ ID NO: 5
      244 <211> LENGTH: 19
      245 <212> TYPE: PRT
      246 <213> ORGANISM: Mycobacterium aurum
      248 <220> FEATURE:
      249 <223> OTHER INFORMATION: SC-S423
      251 <400> SEQUENCE: 5
      252 Asp Arg Phe Trp His Leu Met Asp Glu Pro Gly Pro Leu Ile Glu Thr
      253   1             5             10             15
      255 Ile Glu Tyr
      259 <210> SEQ ID NO: 6
      260 <211> LENGTH: 38
      261 <212> TYPE: DNA
      262 <213> ORGANISM: Artificial Sequence
      264 <220> FEATURE:
      265 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
      267 <220> FEATURE:
      268 <221> NAME/KEY: modified_base
      269 <222> LOCATION: (12)
      270 <223> OTHER INFORMATION: i
      272 <220> FEATURE:
      273 <221> NAME/KEY: modified_base
      274 <222> LOCATION: (18)

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/067,291

DATE: 02/25/2002

TIME: 11:54:03

Input Set : N:\Crf3\RULE60\10067291.txt

Output Set: N:\CRF3\02252002\J067291.raw

275 <223> OTHER INFORMATION: i  
 277 <400> SEQUENCE: 6  
 W--> 278 ttgagacsg gnttcgngc stcsgaycts acstayac 38  
 281 <210> SEQ ID NO: 7  
 282 <211> LENGTH: 32  
 283 <212> TYPE: DNA  
 284 <213> ORGANISM: Artificial Sequence  
 286 <220> FEATURE:  
 287 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer  
 289 <400> SEQUENCE: 7  
 290 ccsggctcgt ccatsagrtg ccagaascgr tc 32  
 293 <210> SEQ ID NO: 8  
 294 <211> LENGTH: 17  
 295 <212> TYPE: DNA  
 296 <213> ORGANISM: Artificial Sequence  
 298 <220> FEATURE:  
 299 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer  
 301 <400> SEQUENCE: 8  
 302 gagccggaag atgttgc 17  
 305 <210> SEQ ID NO: 9  
 306 <211> LENGTH: 24  
 307 <212> TYPE: DNA  
 308 <213> ORGANISM: Artificial Sequence  
 310 <220> FEATURE:  
 311 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer  
 313 <400> SEQUENCE: 9  
 314 ccaccctgcg cgacgtcacc agcc 24  
 317 <210> SEQ ID NO: 10  
 318 <211> LENGTH: 24  
 319 <212> TYPE: DNA  
 320 <213> ORGANISM: Artificial Sequence  
 322 <220> FEATURE:  
 323 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer  
 325 <400> SEQUENCE: 10  
 326 tctacgacgc cgacgagttg atgg 24  
 329 <210> SEQ ID NO: 11  
 330 <211> LENGTH: 32  
 331 <212> TYPE: DNA  
 332 <213> ORGANISM: Artificial Sequence  
 334 <220> FEATURE:  
 335 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer  
 337 <400> SEQUENCE: 11  
 338 tgccatggct gctctttcag acctcggcac ct 32  
 341 <210> SEQ ID NO: 12  
 342 <211> LENGTH: 32  
 343 <212> TYPE: DNA  
 344 <213> ORGANISM: Artificial Sequence  
 346 <220> FEATURE:  
 347 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/067,291

DATE: 02/25/2002

TIME: 11:54:04

Input Set : N:\Crf3\RULE60\10067291.txt

Output Set: N:\CRF3\02252002\J067291.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application Number

L:222 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3

L:278 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6